Doing well and doing good: a multi-dimensional puzzle

Vanina Forget

ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

DOING WELL AND DOING GOOD: A MULTI-DIMENSIONAL PUZZLE

Vanina FORGET

May 2012

Cahier n° 2012-22

(Revised Version 2012-04)

DEPARTEMENT D'ECONOMIE

Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

Doing Well and Doing Good: A Multi-Dimensional Puzzle

Vanina D. FORGET

Ecole Polytechnique ParisTech, Département d'Economie
F-91128 Palaiseau Cedex
France
Email: vanina.forget@polytechnique.edu
Tel: +33 1 69 33 34 17
Fax: + 33 1 69 33 34 27

ABSTRACT

Socially responsible firms are expected by European regulators to create shared value for their shareholders, their stakeholders and society. Yet how to generate profits while providing public good is still academically debated. This paper argues that corporate social responsibility encompasses many diversified policies with variant effects on profitability. Which policy most matters economically? As theoretical guidance lacks, this multidimensional puzzle typically encounters model uncertainty. To account for it, model averaging is presented and applied to a unique database matching the economic and social performances of large European firms. Results support that socially responsible corporate policies do not equally matter to do well and do good. In particular, good business behaviors with customers and suppliers appear core. Strong support is also brought to the co-existence of policies monotonically related to profitability (human resources) and policies with an optimal level (environmental policies). Consequences for business and policy makers are discussed.

RESUME

Pour les régulateurs européens, les entreprises socialement responsables créent de la valeur pour leurs actionnaires, leurs parties prenantes et la société. Toutefois, comment générer du profit tout en fournissant du bien public est débattu académiquement. Cet article soutient que la responsabilité sociale des entreprises englobe des politiques diverses aux effets variables. Laquelle compte le plus économiquement ? Ce puzzle multidimensionnel se heurte en l’absence de théorie unifiée à une incertitude de modèle. Pour la traiter, la moyennisation de modèles est appliquée à une base de données originale associant performances économique et sociale des grandes entreprises européennes. Les résultats montrent que les politiques de responsabilité sociale n’équivalent pas pour réussir économiquement et socialement. Les bonnes pratiques commerciales envers les clients et fournisseurs apparaissent centrales. Des pratiques monotonement liées au profit (ressources humaines) coexistent avec des politiques présentant un optimum (environnement). Les conséquences pour les entreprises et les pouvoirs publics sont discutées.

KEY WORDS: corporate social responsibility; firm performance; model averaging; model uncertainty.

JEL CODES: C11, C52, M14, L2

I. Introduction

According to the European Commission (2011), a 'socially responsible' firm takes responsibility for its impact on society beyond legal constraints. More precisely, the promoted socially responsible firm should integrate social, environmental, ethical, human rights and consumer concerns into its business operations and core strategy with a double aim: maximizing the creation of shared value for its shareholders, stakeholders and society; and identifying, preventing and mitigating its possible adverse impact. Corporate Social Responsibility (labeled CSR hereafter) hence amounts to firms privately providing public good and reducing their negative externalities. Decades of active academic debate cover the ground between such a definition and Friedman's famous New York Times Magazine article title, back in 1970: *'The social responsibility of business is to increase its profits'* (FRIEDMAN [1970]). Most research developed the argument along which a firm providing public good might neither be sacrificing profits nor, as put by Friedman, doing '*hypocritical window-dressing'*, but rather creating value on the long run.

Indeed, a large corpus of empirical literature tested over the last three decades the link between firm financial performance and social performance (see for literature surveys ORLITZKY *et al.* [2003]; MARGOLIS and WALSH [2003]; PORTNEY [2008]). In the most extensive survey and meta-analysis up-to-date, MARGOLIS *et al.* [2009] conclude that corporate social performance has a small, positive and significant effect on corporate financial performance and that it does not destroy shareholder value. However, the mechanisms behind this relationship are complex and not well understood. Thus many scholars along these authors, such as HORVÁTHOVÁ [2010] and SURROCA *et al.* [2010], underline the need to investigate further how organizations can succeed in both economically "doing well" *and* socially "doing good".

Imagine a firm manager whose task precisely is to succeed in "doing well and doing good", or, put in the European Commission words, to create shared value for its shareholders, stakeholders and society. Hereby her job consists in the selection and implementation of corporate policies both generating profits and providing public good (or reducing a social ill). Such policies might encompass highly diversified actions, for instance: investment in pollution abatement processes; publication of a sustainable development report; non discrimination in the workplace; or donations to local charities. Of course, each policy will differently impact society. CSR is thus in essence a multidimensional construct (CARROLL [1979]; WOOD [1991]).

The few empirical analyses tackling the multidimensional nature of CSR simultaneously estimate the effects of different clusters of policies on firm performance and find that they diverge. HILLMAN and KEIM [2001] distinguish direct stakeholders management, positively related to financial performance, from social issues, negatively linked. BRAMMER *et al.* [2006] observe that environment and local community involvement appear negatively correlated with financial performance, whereas human resources are weakly positively linked. Using panel data and a supermodularity approach, CAVACO and CRIFO [2010] even observe the existence of a complementarity premium on specific CSR dimensions (human resources and business behavior towards customers and suppliers), while other practices are relative substitutes (environment and business behaviors). Hereby, to understand how firms can profit from socially responsible policies, it appears core to consider their specific effects.

This paper argues that it is also essential to understand the *relative* effects of the multiple dimensions encompassed in the CSR concept, that is, to understand the relative importance of those policies for firm economic performance. Among the wide range of socially responsible policies, which most matter economically? This question is crucial for shareholders, but also for firm stakeholders and society as it is likely to impact the type of public good provided by the firm.

Yet such a multidimensional puzzle is not trivial to solve. This paper highlights that it typically encounters several dimensions of model uncertainty, namely theory, data and specification uncertainty (DOPPELHOFER and WEEKS [2009]), which have not been dealt with so far in the literature.

Theory uncertainty arises as theoretical guidance is lacking to predict the relative importance and effects of CSR dimensions for financial performance. Indeed, most theoretical work either analyzes CSR as a whole (e.g. CESPA and CESTONE [2007]; BARON [2009]) or focuses on one specific dimension (e.g. environment in SINCLAIR DESGAGNÉ and GOZLAN [2003]). However, depending on the policies considered, diverging effects are observed (e.g. BRAMMER *et al.* [2006]) and we lack predictions on which effect is likely to dominate. For instance, differentiating products based on environmental attributes might create a market opportunity (PORTER and VAN DER LINDE [1995]) while generating a perquisite for managers who like the accolades of environmentalists, creating agency costs (BARON *et al.* [2009]). BROCK and DURLAUF [2001] refer to this ambiguity as the open-endedness of theories: one causal theory being true does not imply the falsity of another.

Data uncertainty also plagues the relationships between social and financial performances, as a wide range of variables have been used to evaluate the former. Mirroring this uncertainty, MARGOLIS *et al.* [2009] identify nine measurements used by researchers to account for CSR, including charitable contributions, environmental performance, revealed misdeeds or self reported social performance. Missing observations and measurement errors also contribute to data uncertainty.

Finally, specification uncertainty arises as researchers are uncertain about the best empirical specification to explain financial performance. Omitted variable drawbacks have already been pointed out by several authors (MCWILLIAMS and SIEGEL [2000]; JOHNSON *et al.* [2009]). Moreover, whereas most literature investigated a monotonic relationship between social and financial performances, BARNETT and SALOMON [2006] made instead a strong case of a hump-shape relationship, thus unveiling the existence of an optimal level of CSR to be reached.

Ignoring model uncertainty generally results in biased parameter estimates, overconfident standard errors and misleading inference (DOPPELHOFER [2008]). Hereby this paper sets out to tackle model uncertainty in the “doing well” and “doing good” debate and to acknowledge the multidimensional nature of CSR. To do so, it introduces in the literature a formal statistical framework that explicitly accounts for model uncertainty, namely model averaging. Model averaging was designed to specifically address model uncertainty by simultaneously weighing evidence for multiple models depicting alternative working hypotheses (DOPPELHOFER [2008]). It allows researchers to examine all possible models, to weigh each model according to quality, and to provide a probability distribution for each coefficient estimate (EICHER *et al.* [2012]).

This powerful technique has already proven insightful in several research fields hampered by model uncertainty, such as growth economics (SALA-I-MARTIN *et al.* [2004]), macroeconomic forecasts (WRIGHT [2008]), policy evaluation (BROCK and DURLAUF [2007]) and finance (PESARAN *et al.* [2009]). Other fields of applications include trade flows (EICHER *et al.* [2012]), labour economics (TOBIAS and LI [2004]) and health economics (JACKSON *et al.* [2009]).

Various methodologies have been proposed to succeed in implementing model averaging. This paper develops an approach built on information-theoretic model averaging and thick modelling (as proposed by KAPETANIOS *et al.* [2008] and PESARAN *et al.* [2009]). In particular, corrected Akaike Information Criteria model averaging (AICc MA) and Schwarz Bayesian Information Criteria model averaging (SIC MA) are discussed. Beyond their original application to the CSR literature, these methods have the potentiality to be relevant to many empirical economists for their relevancy and their straightforward implementation. They are here applied to a database that matches the economic performance of 461 large European firms over the 1998-2007 period with CSR measures provided by the non-financial rating agency Vigeo. This original data enables the study of a fairly complete range of socially responsible corporate policies, namely: environmental policy, good business behavior with customers and suppliers, implication in local communities, human resources management and governance.

The introduction of model averaging in the "doing well" and "doing good" debate brings novel and robust results. First, it unveils the composition of profitably-linked CSR. This composition appears heterogeneous, with different CSR dimensions having different importance. In particular, good business behavior with customers and suppliers are shown to remain crucial. Results also provide strong evidence of the coexistence of corporate policies monotonically linked to economic performance (human resources management) and policies with optimal levels (environment), hence reconciling competing theories. Consequences for corporations and policy makers are discussed.

The remainder of the paper is organized as follows. Section 2 details the methodology. Section 3 presents data. Section 4 displays results and discusses main findings. Section 5 concludes.

II. Methodology

This section first presents the basic empirical framework to study the link between CSR and financial performance, before introducing model averaging and thick modelling. Finally, it discusses their inputs and limits.

II.*1* *Basic Empirical Framework*

Following previous literature (ORLITZKY *et al.* [2003]; MARGOLIS *et al.* [2009]), the basic model to estimate the link between firm performance and CSR is as follows:

$$Base\ Model: \ FP_{it} = \alpha + \beta_1 CSR_{it} + \beta_2 RISK_{it} + \beta_3 LEVERAGE_{it} + \beta_4 SIZE_{it} + \beta_5 R\&D_{it} + \gamma_j\ INDUSTRY_j + \delta_k\ COUNTRY_k + \theta_t\ YEAR_t + \varepsilon_{it} \quad (1)$$

with *i* the number of firms and *t* the year of observation.

FP_{it} is the firm financial performance, measured in this paper either by returns on assets (ROA) or returns on capital employed (ROCE). CSR_{it} is a global CSR measure (here the global Vigeo rating). A standard OLS limits the disaggregation of CSR into multiple policies. Indeed, as CSR policies are part of a company strategy, chosen and implemented by the same firm manager, they are likely to show at least dome degree of multicollinearity, which would bias coefficients. Hence a global CSR measure is here considered with OLS.

$RISK_{it}$ is a solvability ratio control that captures the fact that the more stable a firm, the likelier it is to engage in CSR. $SIZE_{it}$ controls for firm size and is measured as the logarithm of net sales. Larger firms are indeed more likely to encounter major environmental hazards (KONAR and COHEN [1997]), to have larger resources devoted to social investments and to be more exposed to social pressure. $R\&D_{it}$ controls for research and development intensity. Based on a large corpus of theoretical literature linking research and development to long term economic performance, WADDOCK and GRAVES [1997] and later on MCWILLIAMS and SIEGEL [2000] highlight its importance as a control variable. Risk, firm size and R&D intensity are expected to have positive estimates.

$LEVERAGE_{it}$ controls for the firm financial leverage (debt-to-equity ratio) and provides a good indicator of management risk tolerance, which can impact decision making and arbitrage between short and long term performances (WADDOCK and GRAVES [1997]). As such it is expected to negatively impact ROCE and ROA.

INDUSTRY dummies (*j*) are introduced because industrial processes, scale savings, associated pollution levels, stakeholders' activism, exposure and financial risks are sector specific (MARGOLIS *et al.* [2009]). COUNTRY (*k*) is also controlled for as regulations, social demand and stakeholders' pressure vary between European countries. YEAR dummies account for the evolution over the studied period of CSR regulation, public awareness and firm involvement. Finally $\boldsymbol{\varepsilon_{it}}$ is the time variant error term of firm *i* at year *t*.

In this model, data is considered cross-sectional whereas firms count in average three observations over the studied period of time. Yet a panel approach taking temporality into account is prevented by data availability, as discussed in section II.3. Both White's test and Breusch-Pagan's test show some heteroscedasticity in the data. Hence ROGERS (1993)'s estimators based on clusters (firms) are used instead of standard OLS to account for dependent and non-identically distributed error terms. Estimates obtained with this method will be compared to those obtained with model averaging, now presented.

II.*2* *Model Averaging and Thick Modelling*

This section presents the chosen theoretical framework of model averaging and thick modelling. It limits itself to discussing the properties relevant to this paper. For a comprehensive discussion, the reader is referred to PESARAN *et al.* [2009] and KAPETANIOS *et al.* [2008].

Let us start with a set of models. This set is denoted $M = \bigcup_{i=1}^{m} M_i$ where M_i is the *i*-th of the *m* models considered. In this paper, the space of models M consists of all the possible subsets of candidate regressors potentially explaining financial performance, including all five CSR dimensions considered. Our interest is a parameter Δ. The Bayesian framework provides a probability distribution for Δ given M and the observed data *D*. The relevant information data set at time *t* is denoted D_t. The probability distribution $pr(\Delta|D_t, M)$ of the parameter of interest over the space of models considered is given by:

$$pr(\Delta|D_t, M) = \sum_{i=1}^{m} pr(\Delta|M_i, D_t) pr(M_i|D_t) \quad (2)$$

where $pr(\Delta|M_i, D_t)$ is the conditional probability distribution of Δ given a model M_i and the data D_t. It can easily be obtained from standard model specific analysis. $pr(M_i|D_t)$ is the posterior probability of Mi, that is the conditional probability of the model M_i being the true model given the data D_t.

In the Bayesian Model Averaging (BAM) framework, weights used to combine the models under consideration are their respective posterior probabilities $pr(M_i|D_t)$. This approach requires specifications of the prior probability of model M_i and has been the focus of a large corpus of literature.

An alternative to BAM consists in approximating the weights $pr(M_i|D_t)$ by information criteria weights, such as Akaike weights or Schwartz weigths. This approach is developed in KAPETANIOS *et al.* [2008] as the information theoretic Model Averaging, building on the influential work of BURNHAM and ANDERSON [2002]. Applications are expanding and include growth economics (WAGNER and HLOUSKOVA [2009]), finance

(HANSEN [2008]), tourism development (WAN and ZHANG [2009]), health economics (CLAESKENS *et al.* [2006]) and environmental economics (LAYTON and LEE [2006]).

A first weighting scheme proposed in the literature and implemented in the paper is based on AKAIKE's information criteria (1973, 1974), known as AIC. AIC is defined as:

$$AIC = 2k - 2\ln(L) \tag{3}$$

AIC has two components: the negative loglikelihood *–ln (L)*, which measures the lack of model fit to the observed data, and a bias correction factor, which increases as a function of the number of model parameters *k*. More technically, this criterion is an extension of the log-likelihood theory and is based on the Kullback–Leibler information, which can be conceptualized as a 'distance' between full reality and a model. Difference in AIC between two models can thus be analyzed as an estimate of the difference between the Kullback–Leibler distance for the two models. For in-depth analysis of AIC's theory, uses and limits, see KONISHI and KITAGAWA [2007].

AIC has been criticized for its propensity at over-fitting models, meaning that it tends to select too many variables. HURVICH and TSAI [1989, 1995] hence introduced the corrected AIC (AICc), which is AIC with a second order correction for small size samples:

$$AIC_c = AIC + (2(k+1)(k+2))/(n-k-2) \tag{4}$$

with k the number of model parameters and n the number of observations. As n increases, AICc converges to AIC and is asymptotically efficient in both regression and times series. For linear regression, AICc has better bias properties than does AIC. BURNHAM and ANDERSON [2004] thus advocate employing AICc regardless of sample size, which is done in this paper. The reader can refer to MCQUARRIE and TSAI [1998] for further comparisons of AIC and AICc with several competitor criteria for linear regression problems.

Difference between information criteria φ_i is used to rank models. Following BURNHAM and ANDERSON [2002], the likelihood of the model M_i given the data is equivalent to $\exp(-\frac{1}{2}\varphi_i)$. Model likelihoods are normalized to sum up to 1 and referred to as Akaike weights. Akaike weight w_i for model M_i writes:

$$w_i = \exp\left(-\varphi_i/2\right) / \sum_{r=1}^{P} \exp\left(-\varphi_r / 2\right) \quad (5)$$

where $\sum_{i=1}^{m} w_i = 1$. w_i can be interpreted as the probability of selecting model i as being the best if analyses were repeated using independent samples from the same population. This paper implements the use of corrected Akaike weights, based on AICc.

Let us now go back to our parameter of interest Δ. Following BURNHAM and ANDERSON [2002], the averaged estimate $\hat{\hat{\Delta}}$ of Δ is provided by

$$\hat{\hat{\Delta}} = \sum_{i=1}^{m} w_i \, \hat{\Delta}_i \quad (6)$$

with Δ_i the parameter of interest in model M_i and $\hat{\Delta}_i$ the estimate of Δ_i in model M_i. Δ unconditional standard error is given by:

$$\widehat{se}\left(\hat{\hat{\Delta}}\right) = \sum_{i=1}^{m} w_i \sqrt{\widehat{var}\left(\hat{\Delta}_i\right) + \left(\hat{\Delta}_i - \hat{\hat{\Delta}}\right)^2} \quad (7)$$

The second weighting scheme considered in this paper is based on the Schwarz Bayesian Information Criteria (SCHWARZ, 1978), further on SIC, defined as:

$$SIC = -2\ln(L) + p\log(n) \quad (8)$$

BALTAGI [2001] points out that SIC is consistent, meaning that as the sample goes to infinity, the probability that it will choose the correct model from a finite number of models goes to 1. A drawback of this property is that in small samples, SIC tends to select underfitting models. Consequently model selection based on SIC tends not to pick up enough variables in the ‘best’ models. Hereby SIC based model averaging tends to bias downwards variable weights.

This paper considers both approaches to information-theoretic model averaging: the corrected Akaike weights model averaging (AICc MA) and the SIC weights model averaging (SIC MA).

Other weighting schemes have been discussed in the model averaging literature. HJORT and CLAESKENS [2003] discussed the Focused Information Criterion. HANSEN [2007] proposed Mallows model averaging (MMA). WAGNER and HLOUSKOVA [2009] compare AIC MA, SIC MA and MMA. They further introduce, for any given weighting scheme, the so-called inclusion weight as the classical counterpart of the Bayesian posterior inclusion probability of a variable.

A last refinement used in this paper is the combination of information-theoretic model averaging with thick modelling. As detailed in PESARAN *et al.* [2009], thick modelling consists in applying model averaging not to all of the models but only to a given number of top performing models. Individual models are here ranked according to the AICc or SIC criteria. The space of models M under consideration for model averaging is thus reduced to the top performing M' space of models (say the top 25%). Thick modelling has been proposed, among others, by GRANGER and JEON [2004]. Applications include STOCK and WATSON [1999]’s in the context of macroeconomic time series and AIOLFI *et al.* [2001]’s on forecasts of excess returns.

In this paper, the Base Model (1) is first estimated using OLS. Then AICc MA is applied as a benchmark to the Base Model set of variables with global CSR. As year, industry and country controls are kept in all models, the model population counts $2^5 = 32$ possible models based on five variables (global CSR, Risk, Leverage, Size and R&D) and all are considered. In a second step, global CSR is disaggregated into five CSR dimensions (Environment, Business behavior towards clients and customers, Community involvement, Human resources and Governance). The new model population counts $2^{4+5} = 512$ models and model averaging combined with thick modelling is done on the top 100 models based on AICc and SIC rankings.

II.*3* *Inputs and Limits of the Approach*

This paper implements thick modelling, AICc model averaging and SIC model averaging to account for model uncertainty. This section discusses the inputs and limits of the approach.

As highlighted by DOPPELHOFER [2008], the use of model averaging was limited until recent developments in computing power and statistical methods. Here, model fitting and subsequent model averaging is done thanks to the R / java *glmulti* package (CALCAGNO and DE MAZANCOURT [2010]), which renders the implementation of those tools highly straightforward.

A second input of the approach deals with the multicollinearity likely to arise when one considers simultaneously many dimensions of CSR, as previously discussed. CALCAGNO and DE MAZANCOURT [2010] found that BIC variable selection successfully distinguishes the effects of variables correlated at 70% (which exceeds the 62% correlation between CSR dimensions observed in data here). The proposed methodology can thus help bypass the multicollinearity issue inherent to CSR dimensions.

Moreover, this paper tests for curvilinear versus monotonic relationships between the different CSR dimensions and financial performance. Indeed, CSR dimensions are here introduced as three-level factors (bellow average / average / above average). The effect of being above sector average on one CSR dimension is thus separately estimated from the effect of being bellow sector average. This paper hence tests for the existence of optimal levels and identifies different specifications for different CSR dimensions.

Limits of the tools presented are essentially given by data availability. Indeed, model averaging and thick modelling need large datasets for estimations to be reliable. To ensure sufficient data availability, it was not here possible to use data as panel data. Hence causality between corporate social and financial performance is beyond the scope of this paper. Similarly, data availability constrained control variables used in the regressions. For instance, advertisement intensity could not be controlled for. The relevancy of model averaging is thus anchored in data availability, completeness and quality and restricted to empirical fields with sufficient observations.

III. Data

Two sources of data are matched in this research. CSR data is provided by the leading European extra-financial rating agency Vigeo and financial data comes from the database Orbis (Bureau Von Dijk). The database obtained is a non cylindrical panel. Firm is the primary stratification level with 1 to 8 observations per firm (3 observations in average) over the 1998 – 2007 time period. The sample contains 1577 observations on 461 large European listed firms (restricted to 622 observations on 207 firms by the availability of data on research and development expenses) belonging to 13 different countries and 13 industrial sectors.

Vigeo measures extra-financial performance and provides firm ratings based on disclosed information, dialog with the firm and international or European reference frameworks. In particular, this paper uses data on five CSR dimensions: *Environment* (integration of environmental issues into corporate policy, product manufacturing, distribution, use and disposal); *Governance* (balanced power within the board of directors, respect of shareholders' rights, remuneration of key executives and directors, audit and internal controls); *Customers and Suppliers* (respect of business integrity, including sustainable and transparent relationships with customers and suppliers); *Community Involvement* (integration of the firm's impacts on local communities and responsible societal behavior) and *Human Resources* (proactive human resources corporate policy, including career development, continuous improvement of labour relations, quality of working condition). Weak multicollinearity between the ratings is assessed by variance inflation factors (VIF) ranging from 1.12 to 2.19 (see Table I).

As a starting point, this paper postulates that all five CSR dimensions equally matter. Consequently, a *Global CSR* rating is calculated as their arithmetic mean as usually done on such data in this literature (e.g. HILLMAN and KEIM [2001]). Secondly, CSR dimensions are considered separately. For the purpose of the paper, their ratings are transformed into three-level categorical factors: *Worst* (worst-in-class firms; 30%); Average (40%); and *Best* (best-in-class firms; 30%).

Financial measures are given in 2005 USD. Financial performance is measured by two accounting-based ratios: return on assets (ROA) and return on capital employed (ROCE). ROA is the operating income divided by total assets. As such, it measures firm efficiency in generating income from its assets and thus indicates firm profitability, financial leverage put aside. ROCE is the net operating profit after tax divided by capital employed. It thus provides shareholders with a comparison of earnings with capital invested in the firm.

Descriptive statistics can be found in Table I. For a complete description of the variables and data, the reader can refer to the Data Appendix.

TABLE I. - DESCRIPTIVE STATISTICS

Variable	Obs.	Mean	SD.	Min.	Max.	VIF
Global CSR rating	1578	2.94	3.04	1.17	4.83	
Human resources rating	1578	3.03	0.91	1.00	5.00	1.59
Best	516	4.12	0.32	4.00	5.00	
Worst	484	1.91	0.27	1.00	2.00	
Corporate governance rating	1578	2.97	0.92	1.00	5.00	1.14
Best	463	4.13	0.34	4.00	5.00	
Worst	498	1.86	0.35	1.00	2.00	
Customers & suppliers rating	1578	3.05	0.89	1.00	5.00	1.59
Best	534	4.10	0.30	4.00	5.00	
Worst	472	1.93	0.26	1.00	2.00	
Community involvement rating	1578	3.07	0.95	1.00	5.00	1.46
Best	556	4.17	0.37	4.00	5.00	
Worst	498	1.93	0.26	1.00	2.00	
Environment rating	1578	3.06	0.91	1.00	5.00	1.65
Best	517	4.17	0.38	4.00	5.00	
Worst	475	1.92	0.28	1.00	2.00	
ROA	1577	7.86	6.81	-21.29	34.80	
ROCE	1566	14.75	11.40	-77.47	98.41	
Risk (Solvency ratio)	1578	35.54	15.86	-12.66	82.72	
Financial Leverage (Debt to equity ratio)	1578	0.79	0.99	-2.85	7.48	
R & D intensity	622	5.04	6.51	0.00	71.55	
Size (Ln(sales))	1578	15.76	1.43	6.77	19.55	

NOTE:
Table I presents the number of observations, mean, standard deviation, minimum and maximum of the variables used. The Variance Inflation Factor (VIF) measures multicollinearity between the ratings of the five dimensions of Corporate Social Responsibility (CSR) taken into account. “Best” (respectively “Worst”) indicates the ratings of the 30% top (bottom) firms above (under) sectoral average.

IV. Results

This section first establishes result robustness by focusing on control estimates before presenting and discussing main findings.

IV.*1* *Results Robustness*

A specific test is needed to evaluate model averaging result robustness and variable weight significance. In order to do so, a permutation test (also called randomization test) is built and conducted. Test results give us a probability equivalent to a p-value. 1000 permutations were used to compute the test, meaning the smallest possible p-value obtained is 0.001.

As a benchmark, Base Model OLS estimates are compared with AICc MA results on Global CSR. Table II and Table III compare OLS results to AICc MA results for respectively ROA and ROCE. For both methods, the same controls are significant (for OLS) and important (for MA). When significant, control parameters estimated by OLS are of expected signs. In both samples and for both financial performance measures the control estimates are (but for the R&D intensity control for ROCE in full sample) in line with previous literature. The inclusion or not of the R&D variable little biases the estimations.

Then CSR is disaggregated into five dimensions. Both AICc MA and SIC MA are used with the five CSR dimensions on top 100 performing models (thick modelling). Averaged estimates are compared across both methods (AICc MA and SIC MA) and both financial performance measures (ROA and ROCE). Analyses are also conducted on both a restricted sample (with R&D control) and the full sample (without R&D control). Tables IV and VI respectively present AICc MA results for ROA and ROCE and can be compared with tables V and VII displaying SIC MA results. As SIC tends in small samples to underfit models, SIC MA results provide insights on which variables most matter to explain financial performance. For both ROA and ROCE and on both samples, AICc MA results for control variables are consistent in terms of estimate signs, estimate values and variable importance (Tables IV and VI). This consistency supports result robustness.

TABLE II. – OLS AND AICC MODEL AVERAGING ESTIMATIONS FOR RETURN ON ASSETS (ROA) WITH GLOBAL CSR PREDICTOR

	OLS (i)						AICc Model Averaging (ii)					
Dependent	Restricted sample (with R&D)			Full sample (without R&D)			Restricted sample (with R&D)			Full sample (without R&D)		
Global CSR	1.20	**	(0.56)	0.29		(0.36)	0.59	++	(0.40)	0.25		(0.13)
Risk	0.20	***	(0.04)	0.14	***	(0.02)	0.21	+++	(0.00)	0.22	+++	(0.00)
Financial leverage	-0.67	*	(0.36)	-0.46	**	(0.22)	-0.09		(0.04)	-0.81	++	(0.16)
Size	0.56		(0.36)	0.25		(0.29)	1.07		(0.10)	0.32		(0.08)
R&D intensity	-0.03		(0.11)			No	0.10		(0.01)			No
R^2	39.45			24.83			No			No		
F-statistic	9.46	***		7.13	***		No			No		
Observations	622			1577			622			1577		
No. firms	207			461			207			461		
No. models	1			1			32			16		

NOTES:
Table II compares the OLS estimations of the base model (1) explaining ROA to the AICc Model averaging results based on the set of variables used in model (1) with a global CSR variable. Two different samples are used in each case: a sample restricted to data with R&D intensity variable available; and the full sample without the R&D variable.
(i) For OLS, figures in brackets are standard errors. P-values are corrected with Roger's estimator. *p<0.10;**p<0.05; ***p<0.01.
(ii) For AICc Model Averaging, estimates are the averaged parameter estimates ($\hat{\hat{\Delta}}$ in equation (6)) produced by model averaging. Figures in brackets are the unconditional variance ($\widehat{se}\left(\hat{\hat{\Delta}}\right)$ in equation (7)). Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

TABLE III. - OLS AND AICC MODEL AVERAGING ESTIMATIONS FOR RETURN ON CAPITAL EMPLOYED (ROCE) WITH GLOBAL CSR PREDICTOR

	OLS (i)						AICc Model Averaging (ii)					
Dependent	Restricted sample (with R&D)			Full sample (without R&D)			Restricted sample (with R&D)			Full sample (without R&D)		
Global CSR	2.11	***	(0.86)	0.60		(0.46)	1.37	+++	(1.03)	0.40		(0.13)
Risk	0.10		(0.06)	0.03		(0.03)	0.05	++	(0.00)	0.15	+++	(0.00)
Financial leverage	-2.56	***	(0.82)	-2.09	***	(0.44)	-2.69	+++	(0.48)	-3.24	+++	(0.47)
Size	0.37		(0.68)	-0.19		(0.35)	0.89		(0.33)	-0.22		(0.13)
R&D intensity	-0.04		(0.16)	No			0.08		(0.02)	No		
R^2	32.29			16.36			No			No		
F-statistic	6.88	***		8.00	***		No			No		
Observations	618			1566			618			1566		
No. firms	206			457			206			457		
No. models	1			1			32			16		

NOTES:
Table III compares the OLS estimations of the base model (1) explaining ROCE to the AICc Model averaging results based on the set of variables used in model (1) with a global CSR variable. Two different samples are used in each case: a sample restricted to data with R&D intensity variable available; and the full sample without the R&D variable.

(i) For OLS, figures in brackets are standard errors. P-values are corrected with Roger's estimator. *p<0.10;**p<0.05; ***p<0.01.

(ii) For AICc MA, estimates are the averaged parameter estimates ($\hat{\hat{\Delta}}$ in equation (6)) produced by model averaging. Figures in brackets are the unconditional variance ($\widehat{se}\left(\hat{\hat{\Delta}}\right)$ in equation (7)). Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

TABLE IV. - AICC MODEL AVERAGING RESULTS ON 100 BEST MODELS EXPLAINING RETURN ON ASSETS (ROA) WITH CSR DIMENSIONS

		Restricted sample with R&D				Restricted sample without R&D				Full sample without R&D			
Dependent :		Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight
Human Resources	Best	0.48	(0.50)	50	0.57++	0.57	(0.57)	55	0.69++	-0.15	(0.09)	44	0.22
	Worst	-0.76	(0.82)	50	0.57++	-1.01	(0.93)	55	0.69++	-0.15	(0.09)	44	0.22
Customers & Suppliers	Best	0.07	(0.07)	43	0.26	0.06	(0.06)	49	0.24	-0.17	(0.31)	55	0.85++
	Worst	-0.33	(0.33)	43	0.26	-0.29	(0.28)	49	0.24	-1.35	(0.58)	55	0.85++
Governance	Best	0.08	(0.03)	40	0.11	0.08	(0.03)	40	0.11	0.03	(0.01)	42	0.10
	Worst	0.03	(0.02)	40	0.11	0.03	(0.01)	40	0.11	0.01	(0.00)	42	0.10
Environment	Best	-0.63	(0.75)	45	0.42	-0.64	(0.77)	45	0.42	-0.33	(0.28)	46	0.32++
	Worst	-0.32	(0.39)	45	0.42	-0.30	(0.37)	45	0.42	-0.21	(0.15)	46	0.32++
Community Involvement	Best	-0.04	(0.02)	38	0.09	-0.05	(0.02)	42	0.10	-0.13	(0.07)	44	0.20
	Worst	-0.04	(0.01)	38	0.09	-0.04	(0.02)	42	0.10	-0.15	(0.08)	44	0.20
Risk		0.21	(0.00)	100	1.00+++	0.21	(0.00)	100	1.00+++	0.22	(0.00)	100	1.00+++
Financial Leverage		-0.11	(0.05)	44	0.28+	-0.13	(0.07)	48	0.32+	-0.87	(0.15)	62	0.94++
Size		1.13	(0.09)	100	1.00+++	1.17	(0.10)	58	1.00+++	0.33	(0.08)	51	0.71+
R&D intensity		0.10	(0.01)	35	0.60	No				No			
Observations		622				622				1577			

NOTES:

Table IV presents AICc Model Averaging results for the 100 best models explaining Return on Assets (ROA) using CSR dimensions. Three different samples are used to ensure results robustness: a sample restricted to data with R&D intensity variable available; the same restricted sample without the R&D intensity variable; and the full sample without the R&D variable.

"*Estimate*" is the averaged parameter estimate ($\hat{\bar{\Delta}}$ in equation (6)) produced by model averaging. "*Uncond. Var*" is the unconditional variance ($\widehat{se}\left(\hat{\bar{\Delta}}\right)$ in equation (7)). "*No. models*" is the number of models in which a variable is present. "*Weight"* refers to Akaike's weights (equation (5)).

Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

TABLE V. - SIC MODEL AVERAGING RESULTS ON 100 BEST MODELS EXPLAINING RETURN ON ASSETS (ROA) WITH CSR DIMENSIONS

Dependent :		Restricted sample with R&D				Full sample without R&D			
		Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight
Human Resources	Best	0.04	(0.01)	41	0.07	-0.00	(0.00)	41	0.01++
	Worst	-0.09	(0.03)	41	0.07	-0.01	(0.00)	41	0.01++
Customers & Suppliers	Best	0.00	(0.00)	32	0.02+	-0.02	(0.01)	44	0.13+++
	Worst	-0.02	(0.00)	32	0.02+	-0.22	(0.15)	44	0.13+++
Governance	Best	0.00	(0.00)	23	0.00	0.00	(0.00)	40	0.00
	Worst	0.00	(0.00)	23	0.00	-0.00	(0.00)	40	0.00
Environment	Best	-0.02	(0.00)	33	0.02++	-0.01	(0.00)	42	0.01++
	Worst	-0.02	(0.00)	33	0.02++	-0.00	(0.00)	42	0.01++
Community	Best	-0.00	(0.00)	26	0.00	-0.00	(0.00)	41	0.01
Involvement	Worst	-0.00	(0.00)	26	0.00	-0.01	(0.00)	41	0.01
Risk		0.21	(0.00)	100	1.00+++	0.22	(0.00)	100	1.00+++
Financial Leverage		-0.02	(0.00)	43	0.06	-0.54	(0.27)	51	0.59+++
Size		1.23	(0.08)	70	1.00+++	0.17	(0.06)	48	0.35++
R&D intensity		0.06	(0.00)	39	0.35+++	No			
Observations		622				1577			

NOTES:
Table V presents SIC Model Averaging results for the 100 best models explaining Return on Assets (ROA) using CSR dimensions. Two different samples are used to ensure results robustness: a sample restricted to data with R&D intensity variable available; and the full sample without the R&D variable.
"*Estimate*" is the averaged parameter estimate ($\hat{\hat{\Delta}}$ in equation (6)) produced by model averaging. "*Uncond. Var*" is the unconditional variance ($\widehat{se}\left(\hat{\hat{\Delta}}\right)$ in equation (7)). "*No. models*" is the number of models in which a variable is present. "*Weight*" refers to Akaike's weights (equation (5)).
Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

TABLE VI. - AICc MODEL AVERAGING RESULTS ON 100 BEST MODELS EXPLAINING RETURN ON CAPITAL EMPLOYED (ROCE) WITH CSR DIMENSIONS

Dependent :		Restricted sample with R&D				Restricted sample without R&D				Full sample without R&D			
		Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight
Human Resources	Best	0.03	(0.08)	34	0.22+	0.03	(0.12)	45	0.28+	0.12	(0.10)	44	0.16
	Worst	-0.39	(0.50)	34	0.22+	-0.53	(0.79)	45	0.28+	-0.10	(0.09)	44	0.16
Customers & Suppliers	Best	0.27	(0.41)	46	0.41+	0.24	(0.38)	47	0.40+	-0.46	(0.91)	51	0.66+
	Worst	-0.84	(1.49)	46	0.41+	-0.80	(1.42)	47	0.40+	-1.76	(2.47)	51	0.66+
Governance	Best	0.12	(0.08)	28	0.11	0.13	(0.09)	38	0.12	-0.05	(0.04)	42	0.10
	Worst	0.09	(0.06)	28	0.11	0.09	(0.06)	38	0.12	0.01	(0.01)	42	0.10
Environment	Best	-0.40	(0.52)	36	0.23	-0.39	(0.50)	40	0.23	-0.88	(1.54)	47	0.40
	Worst	-0.24	(0.02)	36	0.23	-0.24	(0.31)	40	0.23	-0.42	(0.63)	47	0.40
Community Involvement	Best	0.01	(0.04)	31	0.15	0.01	(0.06)	44	0.19	-0.02	(0.03)	43	0.13
	Worst	-0.22	(0.20)	31	0.15	-0.28	(0.31)	44	0.19	-0.11	(0.07)	43	0.13
Risk		0.04	(0.00)	54	0.61+++	0.05	(0.00)	50	0.69+++	0.15	(0.00)	64	0.99+++
Financial Leverage		-2.69	(0.47)	100	1.00+++	-2.66	(0.50)	98	1.00+++	-3.28	(0.47)	100	1.00+++
Size		1.16	(0.25)	80	0.94	1.22	(0.24)	59	0.96	-0.18	(0.10)	47	0.36+
R&D intensity		0.12	(0.02)	44	0.53	No				No			
Observations		618				618				1566			

NOTES:
Table VI presents AICc Model Averaging results for the 100 best models explaining Return on Capital Employed (ROCE) using CSR dimensions. Three different samples are used to ensure results robustness: a sample restricted to data with R&D intensity variable available; the same restricted sample without the R&D intensity variable; and the full sample without the R&D variable.

"*Estimate*" is the averaged parameter estimate ($\hat{\hat{\Delta}}$ in equation (6)) produced by model averaging. "*Uncond. Var*" is the unconditional variance ($\widehat{se}\left(\hat{\hat{\Delta}}\right)$ in equation (7)). "*No. models*" is the number of models in which a variable is present. "*Weight*" refers to Akaike's weights (equation (5)).
Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

Table VII. - SIC MODEL AVERAGING RESULTS ON 100 BEST MODELS EXPLAINING RETURN ON CAPITAL EMPLOYED (ROCE) WITH CSR DIMENSIONS

		Restricted sample with R&D				Full sample without R&D			
Dependent :		Estimate	Uncond. Var.	No. Models	Weight	Estimate	Uncond. Var.	No. Models	Weight
Human Resources	Best	0.01	(0.00)	28	0.02++	-0.00	(0.00)	35	0.00
	Worst	-0.04	(0.00)	28	0.02++	-0.00	(0.00)	35	0.00
Customers & Suppliers	Best	0.03	(0.00)	35	0.04+++	-0.03	(0.00)	40	0.03+
	Worst	-0.09	(0.03)	35	0.04+++	-0.08	(0.03)	40	0.03+
Governance	Best	0.01	(0.00)	20	0.00	-0.00	(0.00)	32	0.00
	Worst	0.00	(0.00)	20	0.00	0.00	(0.00)	32	0.00
Environment	Best	-0.01	(0.00)	26	0.01	-0.01	(0.00)	38	0.01
	Worst	-0.02	(0.00)	26	0.01	-0.02	(0.00)	38	0.01
Community	Best	0.01	(0.00)	26	0.01++	-0.00	(0.00)	34	0.00
Involvement	Worst	-0.03	(0.00)	26	0.01++	-0.00	(0.00)	34	0.00
Risk		0.02	(0.00)	49	0.27+++	0.15	(0.00)	64	0.97+++
Financial Leverage		-2.89	(0.43)	93	0.99+++	-3.23	(0.48)	100	1.00+++
Size		1.22	(0.27)	63	0.91++	-0.03	(0.00)	46	0.07
R&D intensity		0.07	(0.01)	35	0.29	No			
Observations		618				1566			

NOTES:
Table VII presents SIC Model Averaging results for the 100 best models explaining Return on Capital Employed (ROCE) using CSR dimensions. Two different samples are used to ensure results robustness: a sample restricted to data with R&D intensity variable available; and the full sample without the R&D variable.
"*Estimate*" is the averaged parameter estimate ($\hat{\bar{\Delta}}$ in equation (6)) produced by model averaging. "*Uncond. Var*" is the unconditional variance ($\widehat{se}\left(\hat{\bar{\Delta}}\right)$ in equation (7)). "*No. models*" is the number of models in which a variable is present. "*Weight*" refers to Akaike's weights (equation (5)).
Weight significance is obtained by permutation test: + p<0.10; ++ p<0.05; +++ p<0.01.

IV.*2* *CSR Policies Do Not Equally Matter to Do Well and Do Good*

Let us now focus on CSR estimates. Results obtained with the Global CSR measure support the existence of a positive link with financial performance. Indeed, global CSR rating parameter is estimated to be positive and significant at the 5% level for ROA (1.20) and at 1% for ROCE (2.11) with standard OLS on the restricted sample only. The global CSR averaged estimate is also positive (0.25 to 0.59 for ROA in Table II, 0.40 to 1.37 for ROCE in Table III) but CSR importance (as measured by Akaike's weight) only exceeds 0.50 for the restricted sample. However, further results obtained by disaggregating CSR into multiple cluster of policies (dimensions) show that this global positive relationship hides divergent effects.

To explain ROA, CSR dimensions that stand out as important variables with AICc MA on the full sample are Customers and Suppliers, which clearly stands out (weight 0.85, Table IV), weakly followed by Environment (weight 0.2, Table IV). On the restricted sample, important variables are Human resources (weight 0.57 to 0.69, Table IV) and more weakly Environment (weight 0.42, Table IV). With SIC MA, only Human Resources on the restricted sample and Customers and Suppliers on the full sample are not driven to null weight (Table V). To explain ROCE, Customers and Suppliers (weight 0.40 to 0.41, Table VI) stand out as the important variable both on the restricted and the full samples, weekly followed by Environment (weight 0.40, Table VI), particularly on the full sample. Human Resources comes third with a weaker effect than observed for ROA. SIC MA once again proves very selective but Customers and Suppliers dimension is the only CSR dimension kept.

Hereby a main finding of the paper is that all CSR dimensions do not equally matter to do well and do good. A hierarchy clearly stands out between CSR dimensions, robust and consistent across various samples. This hierarchy is dominated by the Customers and Suppliers dimension. The Human Resources and Environment dimensions appear to have a significant but lesser impact. Finally, the Governance and Community Involvement dimensions do not seem to be linked to financial performance. Let us now focus on the Customers and Suppliers dimension.

IV.*3* *Good Business Behaviors with Customers and Suppliers Remain Core*

The Customers and Suppliers CSR dimension relates to respect of customers, in terms of information and product safety; sustainable and transparent relationships with suppliers; and more generally, business integrity. This paper shows that performance along this policy heavily weights in the composition of profitably-linked CSR.

AICc MA and SIC MA both conclude that having a low performance on this CSR dimension is negatively linked with financial performance. However, a high level is positively linked with financial performance on the restricted sample (implying a monotonic relationship) but negatively linked on the full sample (implying a hump-shaped relationship). Differences between the samples are essentially twofold. First, firms of the restricted samples have communicated their R&D expenses, likely implying an increased transparency. Second, firms of the restricted sample have significantly higher global CSR rating than other firms. Little difference is observed in terms of industry, year or country distributions between full and restricted samples.

As early as 1995, JONES [1995] has shown that companies involved in repeated transactions with stakeholders on the basis of trust and cooperation are motivated to be honest, trustworthy and ethical because the returns to such behavior are high. Whereas it seems fairly intuitive that promoting good-business behaviors with suppliers and customers is likely to create value, at least on the long run, few studies actually quantify it. Part of the effect of this CSR dimension might capture the synergies between reputation, advertising and CSR, which have been more studied (e.g. BROWN *et al.* [2006]). Disentangling consumer relationships, supplier relationships and advertisement effects on financial performance would thus be an interesting path for further research.

IV.*4* *Coexistence of CSR Policies With and Without Optimal Level*

The second major finding of this paper is the coexistence of CSR policies monotonically linked to profitability and of CSR policies with optimal levels.

On the one hand, a monotonic relationship is found between the Human Resources variable and financial performance. This dimension here refers to a proactive human resources corporate policy, including career development, continuous improvement of labour relations and quality of working conditions. Being worst-in-class is found to be negatively linked with financial performance whereas being best-in-class in positively related, indicating a monotonic relationship. This finding is in line with previous works showing that human resources policies can help recruiting motivated employees with team work values, securing firm survival and long-term performance (BREKKE and NYBORG [2008]) and reducing costly employee turnover (PORTNEY [2008]). Empirically, similar findings are made by GALBREATH [2006] who studies employee treatment in 38 top Australian firms; JONES and MURRELL [2001] who focus on the stock returns of the 51 firms included in the 'Working mother' list; and BRAMMER *et al.* [2006], who use the stock returns on the UK market.

On the other hand, a curvilinear relationship between financial performance and the Environment dimension is found on all samples, for both financial performance measures and with both methods. The Environment dimension here encompasses the integration of environmental issues into corporate policy, product manufacturing, distribution, use and disposal. Contrary to the Human Resources dimension, the Environment dimension hence presents an optimum level to be reached.

This finding is a major step in the CSR literature as it reconciles divergent studies such as DERWALL *et al.* [2005], who find a positive link between corporate environmental policies and financial performance, and BRAMMER *et al.* [2006]'s, who found a negative link. In line with BARNETT and SALOMON [2006], who made a strong case for a curvilinear relationship, results here point out that the effect of environment policies on corporate performance depends of the level considered. A step further, our findings highlight that this curvilinear versus monotonic relationship depends on the CSR dimension considered. Indeed, the optimum appears fairly specific to environmental issues, as it not found on the Human Resources dimension. Qualifying in more details those relationships by focusing on causality with panel data would appear as a promising extension of this research.

IV.*5* *Implications for Corporations Seeking to Do Well and Do Good and For Public Policies*

Clearly, this paper results do not imply that to succeed in being socially responsible and "doing well" and "doing good" a firm should heavily invest on its business behavior towards its customers and suppliers, improve its human resources policies, cut down its environmental performance and drop all policies regarding its community involvement and governance. Indeed, several limits of this study have been acknowledged, starting with the causality between CSR dimensions and financial performance, which ought to be the focus of further work. Indeed, long-term effects of socially responsible policies and reverse causality are here bypassed. For instance, being proactive in environmental policies might be costly at year *t* and only possible for firms with excess cash-flows. However, this environmental policy might be well communicated to customers, which might then be willing to pay more for the firm product at year *t+1*. Supporting the existence of interactions between CSR dimensions, CAVACO and CRIFO [2010] found complementarities between good business behaviors and proactive human resources policy, but substitutability between the latter and environmental performance.

However, this paper establishes that CSR is a heterogeneous construct: policies encompassed in this wide concept have different effects on corporations' economic performance; some bear optimum levels (environment) while others don't (human resources); some have little impact (governance) while others weight significantly on the accounting ratios (good business behaviors). Hereby CSR does not come as a bundle to be bluntly promoted. Instead, at the corporation level, it calls for a strategic analysis of the firm's business model in order to carefully select the appropriate CSR policy mix, with a special attention to business integrity towards customers and suppliers.

In terms of public policies, results highlight that for CSR to become a mainstay of the Europe 2020 sustainable development strategy, it needs to be detailed at an implementable policy level, as different policies have diverging effects. It also needs to be built with corporations, as they are core actors to identify which dimensions of CSR can penalize their profitability and which are more relevant to create shared value. In particular, these paper findings suggest that policies targeting the supply chain and customers are more likely to be successfully seized by firms as they directly impact profitability. Conversely, since an optimum level has been found for the environmental performance, high levels of performance along this dimension might be harder to reach by voluntary corporate policies.

5. Conclusion

This paper contributes to the question of how corporations can profitably become “socially responsible”, as defined and promoted by the European Commission. It highlights that this issue constitutes a multidimensional puzzle hampered by model uncertainty. To solve it, it introduces model averaging in the literature.

This powerful tool unveils the composition of corporate policies that matter for profitability. Good business behaviors with customers and suppliers appear to dominate this composition. Strong evidence is also provided on the co-existence of policies with optimal level for financial performance and monotonic policies. In particular, a monotonic relationship is found for human resources management and a hump-shaped relationship for environmental policies, supporting the existence of an optimal level of environmental performance to be reached by corporations.

This research opens a new path to better analyze drivers of how firms can do well and do good. Further work taking into account temporality and causality is still needed before providing reliable CSR strategy advice to organizations seeking to profitably adhere to the principles of CSR. However, this paper findings highlight that firms do not necessarily have interest to promote all CSR dimensions to create value for their shareholders. This result raises questions in terms of the value shared with stakeholders and society and the public good effectively provided. Hereby, further research paths could cover the social side of the equation. In particular, little is known about the type and efficiency of the public good privately provided by corporations under the impulse of public policies.

ACKNOWLEDGMENTS

I acknowledge the support of AgroParisTech (France) and of the Chair for Sustainable Finance and Responsible Investment (Toulouse-IDEI and Ecole Polytechnique, France). I am grateful to Vigeo for granting me access to their data on extra-financial rating. I would like to thank Patricia Crifo for valued guidance; Eric Strobl, Guillem Riguail, Vincent Calcagno, participants at the Environment & CSR Seminar of the Ecole Polytechnique for their helpful remarks; and Jurgis Sapijanskas for insightful suggestions. Naturally, I remain responsible for any mistake still present.

DATA APPENDIX

Corporate Social Responsibility Data

Vigeo ratings on five CSR dimensions are used in this paper: environmental policy rating (Environment), corporate governance (Governance), human resources management (human resources), involvement in local communities (Community involvement) and business behaviors towards customers and suppliers (Customers & suppliers).

Vigeo identifies CSR issues by sector and, for each CSR dimension, specific criteria are selected and weighted according to: CSR type and impact on sector stakeholders; stakeholders' impact exposure; and finally sector risks if the impact is not correctly managed.

Vigeo then rates firm performance on CSR dimensions in terms of leadership, implementation and results. A final score is calculated by firm for each dimension on a 0 (minimum) to 100 (maximum) scale.

For firms to be comparable across sectors, firm scores are benchmarked against their sector average score. The resulting rating is provided on a five-level scale: 'worst-in-class' (5%), 'below sector average' (25%), 'in the sector average' (40%), 'proactive' (25%) and 'best-in-class' (5%). For the purpose of this paper, whose methodology requires sufficient observations per category, those ratings are quantified into a three-level scale: worst (bellow sector average; 30%); average (40%); best (above sector average).

As Vigeo systematically rates the DJ Stoxx 600 firms (largest listed European firms), there is no bias selection in data. Academic work based on Vigeo's data is still scarce (CAVACO and CRIFO [2010]) and promising, notably because it allows researchers to study the European market whereas most previous studies focused on the United States market.

Financial Data

Financial performance and control variables data come from the Bureau Von Dijk's Orbis global database, which is sourced from many different providers. All financial measures are given in 2005 United States dollars and observations with unconsolidated accounting data and more than one subsidiary were not kept. Control for outliers is done by winsorizing at the 2% and 98 % levels ROA and ROCE.

In full sample (not restricted to R&D intensity data availability), firms belong to 17 different countries and 14 industrial sectors (see Table A1).

Pearson correlation coefficients can be found in Table A2.

TABLE A1.- DESCRIPTIVE STATISTICS PER COUNTRY GROUP AND INDUSTRY (FULL SAMPLE)

Variable	ROA			ROCE			Global CSR rating		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Austria	12	8.21	2.97	12	14.16	3.82	12	2.49	0.56
Belgium	42	8.82	6.26	44	18.01	13.49	44	2.66	0.55
Denmark	25	10.14	7.62	25	15.68	9.59	25	2.82	0.91
Finland	42	10.67	7.44	42	19.09	12.95	42	3.29	0.66
France	430	6.64	6.52	416	12.87	9.51	416	3.11	0.62
Germany	189	6.80	6.77	191	13.25	10.13	191	3.14	0.61
Greece	9	10.01	7.16	9	14.05	9.14	9	2.11	0.62
Ireland	27	6.90	5.84	27	12.25	7.33	27	2.48	0.76
Italy	68	8.38	6.51	65	16.46	12.55	65	2.66	0.62
Luxembourg	11	5.64	8.35	11	9.71	11.67	11	3.13	0.75
Netherlands	126	6.16	5.85	126	11.91	10.95	126	3.14	0.57
Norway	17	12.65	9.93	18	20.43	14.27	18	3.06	0.45
Portugal	20	6.21	2.26	20	12.50	4.22	20	2.63	0.46
Spain	96	7.17	6.40	94	13.97	15.74	94	2.72	0.62
Sweden	52	8.61	7.41	55	16.86	12.58	55	2.92	0.61
Switzerland	70	9.91	6.07	71	16.08	8.84	71	2.80	0.72
UK	341	9.46	7.16	340	17.07	12.58	340	3.24	0.63
Car Industry	82	4.64	4.04	83	10.86	8.45	83	3.09	0.64
Trade	129	8.34	7.16	129	17.48	13.07	129	3.06	0.61
Consumer goods	287	10.94	6.91	285	18.99	10.96	285	3.04	0.68
Building	67	6.26	4.62	65	14.65	8.10	65	2.97	0.72
Energy	163	7.95	6.11	161	14.23	10.10	161	3.08	0.63
Equipment	84	7.62	5.99	83	15.86	13.52	83	3.04	0.67
Finance	58	7.53	6.00	58	12.47	7.92	58	3.09	0.66
Hotel industry	54	4.75	4.59	54	10.22	6.46	54	3.05	0.61
Agri-food	91	9.01	4.53	91	16.92	7.58	91	3.03	0.66
Intermediate	196	7.76	6.01	196	13.03	8.29	196	3.05	0.68
ITC	148	7.35	10.24	143	13.74	16.31	143	2.99	0.67
Media	43	7.79	7.26	39	14.47	10.08	39	2.91	0.66
Telecom	76	5.09	7.57	80	9.31	17.39	80	3.02	0.64
Transport	99	5.98	4.33	99	12.21	6.86	99	3.07	0.72

TABLE A2.- PEARSON CORRELATION COEFFICIENTS

	CSR	Human Resources	Governance	Customers & Suppliers	Community Involvement	Environment	ROA	ROCE	Risk.	Financial leverage	R&D intensity	Size
Global CSR	1.00											
Human Resources	0.76	1.00										
Governance	0.56	0.24	1.00									
Customers & Suppliers	0.76	0.49	0.32	1.00								
Community Involvement	0.73	0.45	0.24	0.44	1.00							
Environment	0.76	0.52	0.25	0.50	0.48	1.00						
ROA	0.01	0.01	0.03	0.04	-0.01	-0.01	1.00					
ROCE	-0.01	-0.00	0.05	-0.00	-0.01	-0.05	0.45	1.00				
Risk.	-0.06	-0.03	-0.07	-0.02	-0.08	-0.02	0.37	-0.01	1.00			
Financial leverage	0.00	-0.01	0.02	-0.01	0.02	-0.01	-0.27	-0.11	-0.55	1.00		
R&D intensity	-0.01	0.06	-0.09	0.01	0.02	0.01	0.03	-0.04	0.31	-0.09	1.00	
Size	0.34	0.23	0.16	0.21	0.33	0.28	-0.09	-0.02	-0.33	0.03	-0.22	1.00

REFERENCES

AIOLFI, Marco, Carlo A. FAVERO, and Giorgio PRIMICERI (2001): "Recursive `Thick' Modelling of Excess Returns and Portfolio Allocation", *Working paper of Bocconi University, Milan*. Available at: http://finance.bi.no/~finance/seminar/Favero.pdf. [13]

AKAIKE, Hirotugu (1973): "Information Theory as an Extension of the Maximum Likelihood Principle", in *Second International Symposium on Information Theory,* ed. by B.N. Petrov and F.Csaki, Budapest, Hungary: Akademiai Kiado. [11]

AKAIKE, Hirotugu (1974): "A New Look at the Statistical Model Identification", *IEEE Transactions on Automatic Control,* Vol. 19, p. 716–723. [11]

BALTAGI, Badi H. (2001): *A Companion to Theoretical Econometrics.* Oxford, UK: Blackwell Publishers. [12]

BARNETT, Michael L., and Robert M. SALOMON (2006): "Beyond Dichotomy: The Curvilinear Relationship Between Social Responsibility and Financial Performance", *Strategic Management Journal,* Vol. 27, p. 1101-1156. [6, 28]

BARON, David P. (2009): "A Positive Theory of Moral Management, Social Pressure, and Corporate Social Performance", *Journal of Economics and Management Strategy,* Vol. 18, p. 7-43. [5]

BARON, David P., Maretno A. HARJOTO, and Hoke JO (2009): "The Economics and Politics of Corporate Social Performance", *Working paper of Stanford University Graduate School of Business* N. 1993. [5]

BRAMMER, Stephen, Chris BROOKS, and Stephen PAVELIN (2006): "Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures", *Financial Management,* Vol. 35, p. 97–116. [4, 5, 27, 28]

BREKKE, Kjell Arne, and Karine NYBORG (2008): "Attracting Responsible Employees: Green Production as Labor Market Screening »*, Resource and Energy Economics*, Vol. 30, p. 509-526. [27]

BROCK, William A., and Steven N. DURLAUF (2001): "What Have We Learned From a Decade of Empirical Research on Growth? Growth Empirics and Reality", *The World Bank Economic Review,* Vol. 15, p. 229-272. [5]

BROCK, William A., and Steven N. DURLAUF (2007): "Model Uncertainty and Policy Evaluation: Some Theory and Empirics", *Journal of Econometrics,* Vol. 136, p. 629–664. [6]

BROWN, William O., Eric HELLAND, and Janet Kiholm SMITH (2006): "Corporate Philanthropic Practices", *Journal of Corporate Finance,* Vol. 12, p. 855–877. [27]

BURNHAM, Kenneth P., and David R. ANDERSON (2002): *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach,* Springer-Verlag: New York (2d ed.). [10, 12, 12]

BURNHAM, Kenneth P., and David R. ANDERSON (2004): "Multimodel Inference: Understanding AIC and BIC in Model Selection", *Sociological Methods Research,* Vol. 33, p. 261-304. [11]

CALCAGNO, Vincent, and Claire DE MAZANCOURT (2010) : "Glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models", *Journal of Statistical Software,* Vol. 34, p. 1-29. [14, 14]

CARROLL, Archie B. (1979): “A Three-Dimensional Conceptual Model of Corporate Social Performance”, *Academy of Management Review*, Vol. 4, p. 497–505. [4]

CAVACO, Sandra, and Patricia CRIFO (2010): “The CSR-Firm Performance Missing Link: Complementarity Between Environmental, Social and Business Behavior Criteria?”, *Working Paper HAL-00504747*. Available at: http://hal.archives-ouvertes.fr/hal-00504747/. [4, 29, 32]

CLAESKENS, Gerda, Christophe CROUX, and Johan VAN KERCKHOVEN (2006): “Variable Selection for Logistic Regression Using a Prediction Focused Information Criterion”, *Biometrics,* Vol. 62, p. 972-979. [11]

CESPA, Giovanni, and Giacinta CESTONE (2007): “Corporate Social Responsibility and Managerial Entrenchment”, *Journal of Economic Management and Strategy,* Vol. 16, p. 741-771. [5]

DERWALL, Jeroen, Nadja GUENSTER, Rob BAUER, and Kees KOEDIJK (2005): “The Eco-Efficiency Premium Puzzle”, *Financial Analysts Journal,* Vol. 61, p. 51–63. [28]

DOPPELHOFER, Gernot (2008): “Model Averaging”, in *The New Palgrave Dictionary of Economics,* ed. by L. Blume and S. Durlauf UK: Palgrave Macmillan: London (2nd ed). [6, 6, 14]

DOPPELHOFER, Gernot, and Melvyn WEEKS (2009): “Jointness of Growth Determinants”, *Journal of Applied Econometrics,* Vol. 24, p. 209-244. [5]

EICHER, Theo S., Christian HENN, and Chris PAPAGEORGIOU (2012): “Trade Creation and Diversion Revisited: Accounting for Model Uncertainty and Natural Trading Partner Effects”, *Journal of Applied Econometrics,* Vol. 27, p. 296-321. [6, 6]

EUROPEAN COMMISSION (2011): “A Renewed EU Strategy 2011-14 for Corporate Social Responsibility”, *Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Commitee Of The Regions*, 25 Oct. 2011, p.6. [3]

FRIEDMAN, Milton (1970): “The Social Responsibility of Business Is to Increase its Profits”, *New York Times Magazine,* Sept. 13, p. 32–33. [3]

GALBREATH, Jeremy (2006): “Does Primary Stakeholder Management Positively Affect the Bottom Line?: Some Evidence from Australia”, *Management Decision,* Vol. 44, p**.** 1106-1121. [27]

GRANGER, Clive W. L, and Yongil JEON (2004): “Thick Modelling”, *Economic Modelling,* Vol. 21, p. 323-343. [13]

HANSEN, Bruce E. (2007): “Least Squares Model Averaging”, *Econometrica*, Vol. 75, p. 1175–1189. [13]

HANSEN, Bruce E. (2008): “Least-Squares Forecast Averaging”, *Journal of Econometrics,* Vol. 146, p. 342-350. [11]

HILLMAN, Amy J., and Gerald D. KEIM. (2001): “Shareholder Value, Stakeholder Management, and Social Issues: What’s the Bottom Line?”, *Strategic Management Journal,* Vol. 22, p. 125–139. [4, 16]

HJORT, Nils Lid, and Gerda CLAESKENS (2003): “Frequentist Model Average Estimators”, *Journal of the American Statistical Association,* Vol. 98, p. 879-899. [13]

HORVÁTHOVÁ, Eva (2010): "Does Environmental Performance Affect Financial Performance? A Meta-Analysis", *Ecological Economics*, Vol. 70, p. 52–59. [3]

HURVICH, Clifford M., and Chih-Ling TSAI (1989): "Regression and Time Series Model Selection in Small Samples", *Biometrika*, Vol. 76, p. 297-307. [11]

HURVICH, Clifford M., and Chih-Ling TSAI (1995): "Model Selection for Extended Quasi-Likelihood Models in Small Samples", *Biometrics*, Vol. 51, p. 1077-1084. [11]

JACKSON, Christopher H., Simon G. THOMPSON, and Linda D. SHARPLES (2009): "Accounting for Uncertainty in Health Economic Decision Models by Using Model Averaging", *Journal of the Royal Statistical Society,* Vol. 172, p. 383–404. [6]

JOHNSON, Shane. A., Theodore C. MOORMAN, and Sorin M. SORESCU (2009): "A Reexamination of Corporate Governance and Equity Prices", *Review of Financial Studies*, Vol. 22, p. 4753-4786. [6]

JONES, Thomas M. (1995): "Instrumental Stakeholder Theory: A Synthesis of Ethics and Economics", *Academy of Management Review,* Vol. 20, p. 404–437. [26]

JONES, Ray, and Audrey J. MURREL (2001): "Signaling Positive Corporate Social Performance: An Event Study of Family-Friendly Firms", *Business Society,* Vol. 40, p. 59-78. [27]

KAPETANIOS, George, Vincent LABHARD, and Simon PRICE (2008): "Forecasting Using Bayesian and Information-Theoretic Model Averaging", *Journal of Business and Economic Statistics*, Vol. 26, p. 33-41. [7, 10, 10]

KONAR, Shameek, and Marc A. COHEN (1997): "Information as Regulation: The Effect of Community Right to Know Laws on Toxic Emissions", *Journal of Environmental Economics and Management*, Vol. 32, p. 109–124. [8]

KONISHI, Sadanori, and Genshiro KITAGAWA (2007): *Information Criteria and Statistical Modelling*, Springer-Verlag: New York. [11]

LAYTON, David F., and S. Todd LEE (2006): "Embracing Model Uncertainty: Strategies for Response Pooling and Model Averaging", *Environmental and Resource Economics,* Vol. 34, p. 51–85. [11]

MARGOLIS, Joshua D., and James P. WALSH (2003): "Misery Loves Companies: Rethinking Social Initiatives by Business", *Administrative Science Quarterly,* Vol. 48, p. 268-305. [3]

MARGOLIS, Joshua D., Hillary Anger ELFENBEIN, AND James P. WALSH (2009): "Does It Pay to Be Good... and Does It Matter? A Meta-Analysis and Redirection of Research on Corporate Social and Financial Performance", *Working paper, Harvard University.* Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1866371. [3, 5, 8, 9]

MCWILLIAMS, Abagail, and Donald SIEGEL (2000): "Corporate Social Responsibility and Financial Performance: Correlation or Misspecification?", *Strategic Management Journal*, Vol. 21, p. 603–609. [6, 9]

MCQUARRIE, Alan D. R, AND Chih-Ling TSAI (1998): *Regression and Time Series Model Selection,* Singapore, Malaysia: World Scientific Publishing Company. [11]

ORLITZKY, Marc, Frank L. SCHMIDT, and Sara L. RYNES (2003): "Corporate Social and Financial Performance: A Meta-Analysis", *Organization Studies,* Vol. 24, p. 403-441. [3, 8]

PESARAN, M. Hashem, Christophe SCHLEICHER, and Paolo ZAFFARONI (2009): "Model Averaging in Risk Management with an Application to Futures Markets", *Journal of Empirical Finance,* Vol. 16, p. 280-305. [7, 10, 13]

PORTER, Michael E., and Claas VAN DER LINDE (1995): "Toward a New Conception of the Environment-Competitiveness Relationship", *Journal of Economic Perspectives,* Vol. 9, p. 97-118. [5]

PORTNEY, Paul R. (2008): "The (Not So) New Corporate Social Responsibility: An Empirical Perspective", *Review of Environmental Economics and Policy,* Vol. 2, p. 261-275. [3, 27]

ROGERS, William H. (1993): "Regression Standard Errors in Clustered Samples", *Stata Technical Bulletin,* Vol. 13, p. 19–23. [9]

SALA-I-MARTIN, Xavier, Gernot DOPPELHOFER, and Ronald I. MILLER (2004): "Determinants of Economic Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *American Economic Review,* Vol. 94, p. 813-835. [6]

SCHWARZ, Gideon (1978): "Estimating the Dimension of a Model", *Annals of Statistics,* Vol. 6, p. 461-464. [12]

SINCLAIR-DESGAGNÉ, Bernard, and Estelle GOZLAN (2003): "A Theory of Environmental Risk Disclosure", *Journal of Environmental Economics and Management*, Vol. 45, p. 377-393. [5]

STOCK, James H., and Mark W. WATSON (1999): "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series", in *Cointegration, Causality and Forecasting,* ed. by R. Engle and H. White. Oxford, UK: Oxford University Press. [13]

SURROCA, Jordi, Josep A. TRIBÓ, and Sandra WADDOCK (2010): "Corporate Responsibility and Financial Performance: The Role of Intangible Resources", *Strategic Management Journal*, Vol. 31, p. 463-490. [3]

TOBIAS, Justin L., and Mingliang LI (2004): "Returns to Schooling and Bayesian Model Averaging: A Union of Two Literatures", *Journal of Economic Surveys,* Vol. 18, p. 153-180. [6]

WADDOCK, Sandra A., and Samuel B. GRAVES (1997): "The Corporate Social Performance-Financial Performance Link", *Strategic Management Journal,* Vol. 18, p. 303–319. [8, 9]

WAGNER, Martin, and Jaroslava HLOUSKOVA (2009): "Growth Regressions, Principal Components and Frequentist Model Averaging", *Working paper Institute for Advanced Studies, Economics Series 236, Wien, Austria*. Available at: http://www.eea-esem.com/files/papers/EEA-ESEM/2009/827/grpcma4.pdf. [10, 13]

WAN, Alan T. K., and Xinyu ZHANG (2009): "On the Use of Model Averaging in Tourism Research", *Annals of Tourism Research,* Vol. 36, p. 525–532. [11]

WOOD, Donna J. (1991): "Corporate Social Performance Revisited", *Academy of Management Review,* Vol. 16, p. 691–718. [4]

WRIGHT, Jonathan H. (2008): "Bayesian Model Averaging and Exchange Rate Forecasts", *Journal of Econometrics,* Vol. 146, p. 329-341. [6]